William & Mary
W&M ScholarWorks

Undergraduate Honors Theses | Theses, Dissertations, & Master Projects

4-2015

Spread of Adaptation and Migration Load in a Spatially Structured Population

Scott W. Nordstrom
College of William and Mary

Follow this and additional works at: https://scholarworks.wm.edu/honorstheses

 Part of the Evolution Commons, Other Applied Mathematics Commons, Population Biology Commons, and the Toxicology Commons

Recommended Citation
Nordstrom, Scott W., "Spread of Adaptation and Migration Load in a Spatially Structured Population" (2015). *Undergraduate Honors Theses*. Paper 188.
https://scholarworks.wm.edu/honorstheses/188

This Honors Thesis is brought to you for free and open access by the Theses, Dissertations, & Master Projects at W&M ScholarWorks. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.

Spread of Adaptation and Migration Load in a Spatially Structured Population

A thesis submitted in partial fulfillment of the requirement
for the degree of Bachelor of Science in Biology from
The College of William and Mary

by

Scott Nordstrom

Accepted for Honors

M. Drew LaMar, advisor

Dan Cristol

Leah Shaw

John Swaddle

Williamsburg, VA
April 30, 2015

Acknowledgements

Thank you to Professors Dan Cristol and John Swaddle for the project idea, as well as for consultation and help with background information throughout the exploratory phase of this project. Thank you to Natalie, Michael, Derin and Charlie for tolerating my occasional anxiety and frequent, prolonged disappearences. Greatest of thanks to Professor Drew Lamar, who served as a vital resource for mathematical, biological and computational, information and advice, as well as for serving as a valuable editor. His patience and helpfulness made this process logical and sensical.

Abstract

Mercury contamination reduces fledging probability in birds. Mercury has been introduced to the South River in the Shenandoah, creating differences in habitat quality on a landscape of fragmented forest patches. To study the possible outcomes of the spread of this adaptation through populations in the Shenandoah, we construct and implement a mathematical model that features common life history traits, including dispersal and nest competition, of a generic bird species. To see which processes or parameters have the largest effect on allele frequencies, population sizes and reproductive output (a proxy for fitness), we use partial rank correlation. We conclude that in a simple two-patch instance of the model, local selection pressures and asymmetry of migration have the largest effects on allele frequencies and fitnesses. In half of cases observed, the mercury tolerance allele did not establish in either patch, but it did reach fixation in both patches in 20% of trials run. We also find that in a two-patch instance, allele frequencies in the contaminated and uncontaminated patches tend to correlate with each other very well, suggesting that there is a homogenizing process which couples allele frequencies.

Contents

List of Figures

List of Tables

List of Variables and Parameters

	Description of variable	Introduced In
t	Time step	
$S_{i,t}$	Subadults (born in $t-1$) in population i at end of time t	2.1.1
$A_{i,t}$	Adults (born in $t-2$ or earlier) in population i at end of time t	2.1.1
ϕ_i	Total nest sites in habitat i	2.1.2
ϕ'_i	Vacant nest sites in i	2.1.3
$S^b_{i,t}$	Breeding subadults in i at end of time t	2.1.2
$A^b_{i,t}$	Breeding adults	2.1.2
$S^f_{i,t}$	Floating subadults	2.1.2
$A^f_{i,t}$	Floating adults in i	2.1.2
$A^n_{i,t}$	Breeding adults in $t-1$ who retain nest at beginning of time t	2.1.3
$a^S_{i,t}$	Frequency of tolerant allele among subadults in i at end of t	2.1.6
$a^A_{i,t}$	Frequency of tolerant allele among adults in i	2.1.6
$\hat{A}^b_{i,t}$	Post-dispersal number of breeding adults in patch i during time t	2.2.1
$\hat{A}^f_{i,t}$	Post-dispersal number of floater adults	2.2.1
$\hat{S}^b_{i,t}$	Post-dispersal number of breeding subadults	2.2.1
$\hat{S}^f_{i,t}$	Post-dispersal number of floater subadults	2.2.1
$g^y_{t,i}$	Vector of genotypic frequencies for stage class y in patch i	2.1.4
$g'^y_{t,i}$	Vector of tolerant genotypic frequencies, stage class y in patch i	2.1.4
	Description of parameter	
f_S	Eggs laid per subadult	
f_A	Eggs laid per adult	
σ_H	Probability of hatchling surviving to subadult (surviving first year)	
σ_S	Probability of subadult surviving to adult (surviving second year)	
σ_A	Probability of adult surviving (surviving third and subsequent years)	
l_{tol}	Fledging probability for tolerant homozygote hatchling	
l_{wild}	Fledging probability for wild type homozygote hatchling	
$\mathbf{L}_S$	Weighted fledging probability for breeding subadults	2.1.6
$\mathbf{L}_A$	Weighted fledging probability for breeding adults	2.1.6
$\mathbf{b}_S$	Per capita birth term for subadults	eq 2.3
$\mathbf{b}_A$	Per capita birth term for adults	eq 2.3
α	Exponential decay rate of dispersal likelihood	
ι	Proportion of nesters guaranteed nests in next time step	
d_{ij}	Effective distance from patch i to patch j	
d	Breeding efficiency when all breeding territories full	

Chapter 1

Introduction

1.1 Motivation

Ecological processes do not always happen over homogeneous landscapes; this invalidates a key simplifying assumption in many spatial models. Ecological metapopulation theory has provided a way to simply and elegantly study population dynamics on heterogeneous landscapes [17]. A metapopulation is a collection of discrete interacting subpopulations, arranged spatially within a network of habitat patches [18].

Between 1929 and 1950, an industrial plant in Waynesboro, Virginia released mercury into the South River and South Fork Shenandoah River in the Shenandoah Valley [12]. It has been demonstrated that this mercury has infiltrated terrestrial food webs as birds consume aquatic insects [10]. Due to biomagnification, birds are showing hazardous levels of mercury in their blood. Presence of mercury has been demonstrated in laboratory settings to reduce the reproductive output of exposed birds, mainly by decreasing the proportion of hatched birds that fledge from the nest [38]. However, there is evidence that there is heritable variation in avian response to mercury [37]. Studies of killifish have demonstrated heritable antioxidant defense to aquatic contaminants [25] [26], suggesting that such a mechanism might exist for mercury tolerance in birds.

Mercury contamination is mostly confined to the affected rivers and their flood plains, but birds are highly mobile and move between patches of forest habitat. Metapopulation theory has been used to study birds [14][28][33]. Bird populations on fragmented woodland [41] and urban [29] landscapes have been demonstrated to have characteristics of metapopulations, such as genetically distinct subpopulations and recolonization events. Scrub jays populations, for example, show higher levels of genetic differentiation as the size of gaps between habitats grows [9]. Tits in mixed deciduous and evergreen habitats demonstrate adaptation to the more

common forest type, yet still colonize less suitable forest patches where they are maladaptive and have lower reproductive success, creating source-sink dynamics [11]. We will assume for this project that the metapopulation theory approach is applicable to bird populations in the fragmented habitat patches of the Shenandoah, i.e., that birds breed in and disperse between distinct subpopulations.

Widespread tolerance to mercury contamination could be an adaptation, created by local selection on tolerance genes. However, surrounding populations in uncontaminated areas may flood the contaminated area with locally maladaptive wild-type alleles, establishing polymorphism and reducing overall fitness. The weaker the polymorphism is (i.e., the more prevalent adaptive tolerance is), the greater the risk of removing legacy contamination, as removing mercury would reverse selection pressure by favoring maladaptive wild-type alleles that are in a minority. It is important, then, to assess level of adaptation of populations directly exposed to mercury when considering river restoration, as a fully adapted population would suffer from absence of the contaminant. A spatially structured model can be used to study the effects of local immigration on the level of adaptation in contaminated areas.

However, the effects of contamination do not have to be limited to areas where the contaminant is present. Theodorakis noted that radionuclide-stressed kangaroo rat populations and reference populations did not show significant genetic variation unless controlling for migration [35]. Spromberg et. al. were able to simulate distant effects of a contaminant on a metapopulation, paying special attention to patch arrangement [34]. This means that the genetic effects of contamination may be felt outside of contaminated areas; when this causes a reduction in fitness, it is called a migration load. Dispersal of a locally adapted allele such as mercury tolerance into areas where it is not beneficial may present problems if breeding into local populations occurs and produces maladapted offspring. One such example of this comes from fisheries management, where escaped captive-bred fish are maladapted for the wild, and have lowered reproductive output and less fit offspring than wild fish [2]. It is important, then, not just to study prevalence of tolerant alleles in contaminated areas but also in uncontaminated areas, which may see a reduction in overall fitness despite a lack of direct exposure.

Such phenomena can be studied using mathematical modeling and simulations. For this project, we created a model of a bird population occupying discrete habitat patches. Bird population densities and frequency of a tolerance allele are recorded in each patch over discrete time steps. We used latin hypercube sampling to run the model over 1500 parameter combinations for a two-patch instance, allowing variation in life history traits such as fecundity, behavioral traits such as dispersal likelihood, and environmental traits such as inter-patch distance and selection

pressure. We used partial rank correlation to compare model inputs to resulting allele frequencies, population sizes and fitness estimates at equilibrium. This will tell us which parameters have the strongest effects on population sizes and tolerance prevalence, and by proxy also the likelihood of adaptation or migration load.

1.2 Predictions

First, it is natural to assume that a stronger local selection pressure leads to a lower probability of polymorphism. The mechanism here is that natural selection favors birds with the locally adaptive allele, who have more fledgelings and thus more offspring in the next year. Thus, our first prediction is that selection strength against the locally maladaptive allele will have a strong effect on the frequency of the tolerance allele in that patch. For example, stronger selection against the wild type allele in a contaminated patch should lead to a higher frequency of the tolerant allele in that patch.

Second, we note that we are studying a phenomenon that relies on dispersal between patches. Without dispersal, in the absence of genetic drift or random mutations (neither of which we will incorporate into this model) and over dominance, we would expect each subpopulation to reach fixation of its adaptive allele [22]. Movement of maladapted individuals introduces maladapted alleles into each subpopulation. Thus, gene flow is an important process worthy of study in this project. Theory predicts that strong enough gene flow can wipe out or at least compromise local adaptation [3][21][23] (this has also been demonstrated empirically [27]), and conversely that strong local directional selection can minimize effects of gene flow and create population differentiation [6] (although not in all cases [15]). Our second prediction is that lower levels of gene flow will lead to higher frequency of the locally adapted allele, and conversely that higher gene flow will lead to higher prevalence of the locally maladaptive allele.

We suppose that dispersal between two patches is enhanced by two factors: easier travel between patches and higher reproductive rates that create crowding and necessitate moving between patches. In the case of ease of travel between patches, barriers to dispersal [19] or small dispersal radius [9] impede movement and lead to genetically distinct subpopulations. Our third prediction is that parameter inputs associated with easier travel between patches, e.g., larger dispersal radius, shorter inter-patch distances, or more available nesting sites, will lead to higher rates of population mixing and thus increase the prevalence of maladaptive alleles. Another factor that may lead to more gene flow is a larger dispersing population. If population growth persists when all breeding territories are occupied, and territories can be no further compacted to make new breeding sites, then there must be a class

of non-reproductive "floaters" that compete for nesting sites [31]. In the quest for nests, these floaters may move among patches, even if into suboptimal habitats [7]. Our fourth prediction is that increasing fecundity and survivorship or decreasing strength of density dependence will increase local maladaptation by creating more floaters who will breed into non-optimal subpopulations.

The gene flow question can be viewed a different way: symmetry of dispersal. Models have demonstrated that asymmetrical gene flow between patches with opposing selection pressures lowers the viability of the whole population [5][39]. Theory also shows that asymmetrical migration lowers the likelihood of polymorphism in the lower-immigration patch [21] (this has also been demonstrated empirically [4]). Our fifth and final prediction is that increasing asymmetry between contaminated and uncontaminated patches in the model will increase local maladaptation. This can be studied by modifying the number of breeding territories in a patch, or by increasing the relative selection differential in a patch, which can create a "migrational meltdown" wherein maladaptation opens up more breeding territories, which in turn attracts more maladapted individuals [21].

Chapter 2

Model Development

This model has discrete time steps, each simulating the passing of one breeding season. A time step begins with settlement of open breeding territories. Surviving breeders and floaters from the previous year compete for these nesting sites, and in the process the floaters may disperse among patches. When nests are settled, breeders lay eggs which hatch and fledge. After fledging, all birds either mature or die; those birds that die open breeding territories for settling at the beginning of the next time step. This is process illustrated in figure 2.1. All processes are deterministic.

2.1 Life History Dynamics

This section developes within-subpopulation dynamics and ignores interactions between subpopulations; this is similar to describing a single subpopulation in isolation. Because there is only one subpopulation to describe, notation in this section will not include subscripts for patch numbers.

2.1.1 Stage Classes and Transitions

We model a bird that has two stage classes (Cristol, Fovargue, unpublished), a subadult stage class S (one year old) and an adult stage class A (older than one year). Both classes are reproductive, although they give birth at different rates and exhibit different survival probabilities. In a time step, the breeding subadults S^b and adults A^b will give birth to hatchlings, which develop in that same time step into subadults. After giving birth, subadults and adults will mature or die, with surviving subadults moving into the adult class with surviving adults. This means that all individuals that breed in time t and survive to $t+1$ must be adults in $t+1$.

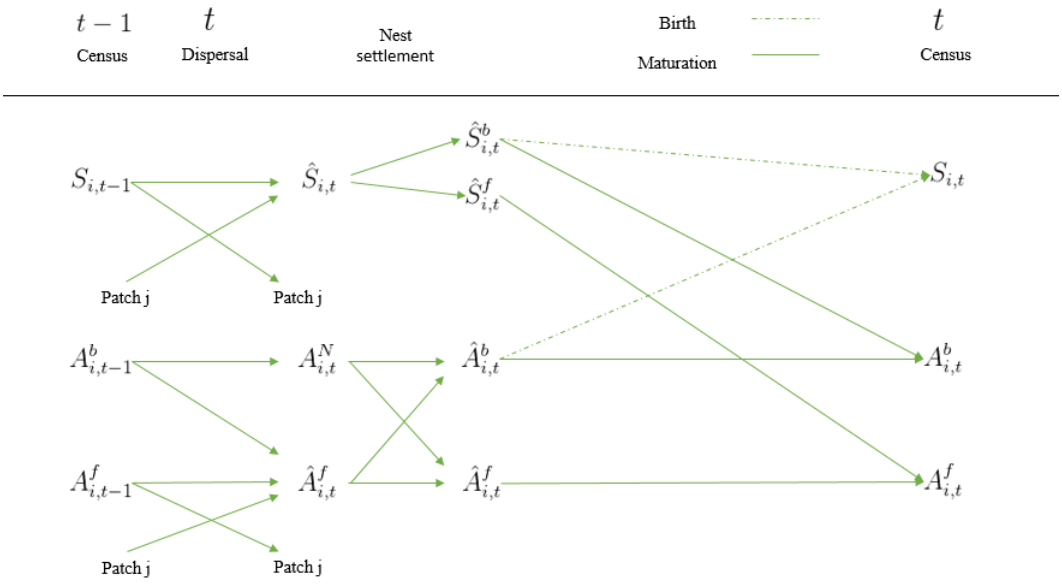


Figure 2.1: Life history diagram showing events during one time step. Time step begins between the $t-1$ census and t dispersal.

Population densities are recorded after reproduction and maturation. By our timing, this means that densities are recorded at the very end of each time step. That is to say, S_t and A_t represent both population densities at the end of time step t and at the very beginning of time step $t+1$. Simplistically, the population's growth can be defined by the following equations:

$$S_t = \mathbf{b}_S S^b_{t-1} + \mathbf{b}_A A^b_{t-1} \tag{2.1}$$

$$A_t = \sigma_S S_{t-1} + \sigma_A A_{t-1} \tag{2.2}$$

$\mathbf{b}_S$ and $\mathbf{b}_A$ are functions that depend both on population size and allele frequencies. σ_S and σ_A are constants on the range $[0, 1]$.

In the birth functions, subadults and adults have different fecundities, f_S and f_A. Eggs laid have different probabilities of surviving to fledge; survivorship is dependent on genotypes and the presence/absence of contaminant. For now, we define $\mathbf{L}_S$ and $\mathbf{L}_A$ as functions for fledging probability in subadults and adults, respectively; they will be defined completely in 2.1.6. Finally, there is the probability of surviving the winter from fledging to subadulthood (σ_H). In the absence of density-dependent regulation, the number of offspring per subadult would be $f_S \mathbf{L}_S \sigma_H$, and the number of offspring per adult would be $f_A \mathbf{L}_A \sigma_H$.

We assume that our bird species has a non-reproducing floater class, created by

population growth that persists when the population size exceeds the number of available breeding territories [31]. As the size of the floater class increases, resource and habitat strain cause population growth to decrease. There are two commonly-employed equations to model discrete-time density dependence, the Beverton-Holt equation and the Ricker equation. We employ the Beverton-Holt equation because we assume that the rate that hatchlings mature into subadults is constant (one breeding season), and that there is no time lag between vital rates and environmental conditions [36]. One advantage to the Beverton-Holt equation is that it is easier to solve analytically for a steady-state; this becomes useful in determining initial conditions for simulations (see 3.3). In this equation, reproductive output continues to decrease as more floaters enter the population (i.e., the density is dependent on the total number of birds, not just the number of breeders). This assumption is justified by the fact that floaters compete with reproducing birds for food and breeding territories, stressing reproducing birds and their offspring. The Beverton-Holt equation, which does not have an Allee effect, models density dependent growth and takes the form

$$\text{per capita growth} = \frac{f}{1 + \delta(S_t + A_t)}$$

where f is maximum fecundity (i.e., as $S_t + A_t \to 0$) and δ is a constant. Let $0 \leq d \leq 1$ be the proportion of maximum fecundity a bird achieves when all ϕ breeding territories are occupied. This introduces density dependence, where reproductive capacity continuously and monotonically decreases (unless $d = 1$) as population size increases. We solve for δ in the following equation:

$$\begin{aligned} df &= \frac{f}{1 + \delta\phi} \\ \frac{1}{d} &= 1 + \delta\phi \\ \frac{1-d}{d} &= \delta\phi \\ \delta &= \frac{1-d}{d\phi} \end{aligned}$$

Maximum fecundity is $f_S \mathbf{L}_S \sigma_H$ for subadults and $f_A \mathbf{L}_A \sigma_H$ for adults. Our per-capita birth terms, then, are

$$\mathbf{b}_S = \frac{f_S \mathbf{L}_S \sigma_H}{1 + \frac{1-d}{d\phi}(S_t + A_t)} \qquad \mathbf{b}_A = \frac{f_A \mathbf{L}_A \sigma_H}{1 + \frac{1-d}{d\phi}(S_t + A_t)} \tag{2.3}$$

meaning that the recruitment rate is:

$$S_{t+1} = \mathbf{b}_S S_t^b + \mathbf{b}_A A_t^b = \frac{\sigma_H\left(f_S \mathbf{L}_S S_t^b + f_A \mathbf{L}_A A_t^b\right)}{1 + \frac{1-d}{d\phi}(S_t + A_t)} \tag{2.4}$$

where S_t^b,A_t^b are the breeding subadults and adults in the patch (see next section).

2.1.2 Breeding Birds and Floaters

We assume in this model that if there are open breeding territories, floaters will fill them. Thus, it is impossible to have open breeding territories and floaters in the same patch. Let ϕ be the number of breeding territories in a patch. When $S_t + A_t \leq \phi$, there are no floaters, and the number of breeding birds is equal to the population size (i.e., $S^b = S$ and $A^b = A$). However, when $S_t + A_t > \phi$, some birds do not get nests, and there are floaters [31]. Floater birds move between patches in search of new breeding territories (this process is described in 2.2.1). Since only birds with breeding territories can reproduce, and there are a fixed finite number of breeding territories in a patch, there is a maximum number of breeders the whole population can sustain. This prevents the population from growing arbitrarily large.

At the beginning of each simulation, breeding territories are divided proportionally among subadults and adults. This means that

$$S_0^b = \min\left(1, \frac{\phi}{S+A}\right) S \qquad\qquad A_0^b = \min\left(1, \frac{\phi}{S+A}\right) A$$

For population levels below ϕ, the breeding population makes up the entire population, but for population levels above ϕ, the breeding population sums up to ϕ. The number of floaters in any time step is defined by

$$S_t^f = S_t - S_t^b \qquad\qquad A_t^f = A_t - A_t^b$$

During competition for nests, there is no distinction between "floater" subadults and "breeding" subadults. Since floaters are birds that do not have nests, in the model we can consider all subadults as floaters when nest competition occurs, since subadults are those birds born in the previous time step and thus have not yet possessed a breeding territory. For this reason (and ease of notation), we will drop the f superscript from subadults (i.e., simply write S instead of S^f) when

considering calculations during nest occupancy. S^f will be reserved for subadults that are not able to attain nests *after* nest competition has occurred.

2.1.3 Nest Competition

Let ι be the proportion of surviving breeders at the end of time $t-1$ that are guaranteed a nest in time step t. $1-\iota$, then, is the proportion of occupied nests that breeders get dislodged from. ι will be fixed in each model iteration and does not change over time. Define A_t^n as the number of breeding birds in $t-1$ guaranteed their nest in time step t.

$$A_t^n = \iota(\sigma_S S_{t-1}^b + \sigma_A A_{t-1}^b)$$

At the beginning of time step t, the number of available breeding spots is $\phi_t' = \phi - A_t^n$. Note that for $\iota = 0$, no breeders are guaranteed nests, and $\phi_t' = \phi$, meaning that all nests are subject to competition by all subadults and adults. Assume that the ϕ_t' spots are divided up proportionally among newborn subadults (S_t) and floaters (A_t^f); we do not assume that a bird's competing ability depends on its age. The number of breeding individuals in a time interval, then, is:

$$S_t^b = \min\Big(\frac{\phi'}{S_t + A_t^f}, 1\Big) S_t \tag{2.5}$$

$$A_t^b = \min\Big(\frac{\phi'}{S_t + A_t^f}, 1\Big) A_t^f + A_t^n \tag{2.6}$$

These values for breeders are used in equation 2.4 to calculate the number of subadults in the next time step.

2.1.4 Allele frequencies

We assume the mercury tolerance mechanism is a single-locus gene with only two alleles; mercury tolerance and a wild type. A plausible mechanism for this is a gene for a detoxification protein, which can isolate or remove mercury from the body. Similar examples of gene complexes have been found with killifish in the Elizabeth River [25] [26]. It has been demonstrated that zebra finches have a heritable capacity for tolerance to mercury (a non-decrease in fledging rates); in the absence of mercury, tolerant individuals have lower fledging rate than intolerant individuals [37]. The dominance of this gene will be controlled with a parameter h, controlling the fitness of heterozygotes.

It should be noted here that the tolerance mechanism is not well understood. We assume that it is a single-locus detoxification protein, although this may be a naive assumption. Fledging probability is the only parameter our model uses with observed differential response to mercury exposure. Thus, this model assumes that the fledging probability of birds is a function of their genotypes and the mercury dosage in their environment; all other parameters are constant across genotypes.

Let a_t^S be the frequency of the tolerance allele in the subadult population at the end of time t, and likewise for a_t^A in the adult population (this means that $(1-a_t^S)$ and $(1-a_t^A)$ are the frequencies of the wild type allele). Let g_t^S be a vector of genotypic frequencies among breeding subadults; its entries correspond to tolerant homozygote, heterozygote, and intolerant wild-type, respectively.

$$g_t^S = \begin{bmatrix}(a_t^S)^2 & 2a_t^S(1-a_t^S) & (1-a_t^S)^2\end{bmatrix}^T$$

g_t^A is defined similarly for breeding adults using a_t^A. Note that the l_1 norms of g_t^S and g_t^A are equal to 1; all individuals must have one of these three genotypes, and any calculation using g_t^S accounts for all birds.

2.1.5 Genotype inheritance

For simplicity, in this subsection we will disregard stage class, labeling the frequency of the tolerant allele p and the frequency of the wild-type allele q. Genotypic frequencies, then, are p^2, $2pq$, and q^2. We assume in the model the number of males and females is the same, and that males and females have the same genotypic frequencies. This means that p is the same for males and females (and by extension, so is q). We assume that all mating is random, and that fecundity is not affected by mercury presence or genotype (thus all birds of the same stage class have the same fecundity). Combinatorially, the expression $(p^2+2pq+q^2)^2$, which expands to $p^4+4p^3q+6p^2q^2+4pq^3+q^4$, gives the probability of each individual combination of parental genotypes (the probability of having two homozygotic tolerant parents, for example, is p^4). We will group together all of these probabilities by the expected genotypes of their eggs.

For example, consider all parental combinations which can yield a tolerant homozygote offspring: two tolerant homozygotes, two heterozygotes, or one tolerant homozygote and one heterozygote. Using a Punnett square, one knows that all offspring of two homozygotic tolerant parents will be homozygotic tolerant. One half of expected offspring of a homozygotic tolerant and a heterozygotic pair will be homozygotic tolerant. One quarter of offspring of two heterozygotes will be homozygotic tolerant. The probability of having a homozygotic tolerant offspring, then, is:

$$
\begin{aligned}
&(p^2)^2 + \frac{1}{2}(2)p^2 2pq + \frac{1}{4}(2pq)^2 \\
&= p^4 + 2p^3q + p^2q^2 \\
&= p^2(p^2 + 2pq + q^2) \\
&= p^2
\end{aligned}
$$

This means that the probability of random mating producing a homozygote tolerant egg is p^2, or exactly the proportion of breeders with that genotype. Note that this is merely an example of Hardy-Weinberg equilibrium, as there is no selection acting upon the number of eggs laid. Similarly, it can be shown that the probability of mating producing a heterozygote egg is $2pq$, and the probability of mating producing a homozygote recessive egg is q^2. This justifies using allele frequencies at the end of time step t (those of the parents) to predict the genotypes of eggs laid (those of the offspring) in $t + 1$ before fledging, when selection occurs.

2.1.6 Differential fledging probabilities

Equation 2.4, the recruitment rate of subadults, is dependent on $\mathbf{L}_S$ and $\mathbf{L}_A$. These are functions for the respective probability that an individual subadult or adult fledges, depending on local mercury contamination and genotype. For recruitment, we count the expected number of fledgelings given a certain number of breeders and their genotypes. The expected number of fledgelings in a pool of breeders is the product of the number of fledgelings born to parents of one genotype multiplied by the probability of a parent having that genotype, summed up over each genotype. Thus, $\mathbf{L}_S$ and $\mathbf{L}_A$ are respectively average fledging probabilities of eggs laid by a set of breeding subadults and adults, weighted by allele frequencies. From the laws of probability, this can be stated as:

$$
\begin{aligned}
\Pr(\text{fledge}) = {} & \Pr(\text{fledge}|\text{tolerant homozygote}) \Pr(\text{tolerant homozygote}) \\
& + \Pr(\text{fledge}|\text{heterozygote}) \Pr(\text{heterozygote}) \\
& + \Pr(\text{fledge}|\text{wild type homozygote}) \Pr(\text{wild type homozygote})
\end{aligned}
$$

There is no evidence thus far suggesting that mercury affects the offspring of subadults and adults differently; we then assume that the fledging probabilities for adults and subadults of the same genotype are the same. Define l as a vector of fledging probabilities for different alleles in a given patch. l_{tol} is the fledging probability of a homozygotic tolerant egg, i.e., Pr(fledge|tolerant homozygote), and

l_{wild} the fledging probability for a homozygotic wild type egg,
Pr(fledge|wild type homozygote). Note that these variables are patch-dependent; fledging probability for one genotype is different in a contaminated patch than in an uncontaminated one.

$$l = \begin{bmatrix} l_{tol} & (h)l_{tol} + (1-h)l_{wild} & l_{wild} \end{bmatrix}^T$$

h serves as a heterozygosity modulator; when $h = 1$, the tolerant allele is dominant, when $h = 0$, the tolerant allele is recessive. Intermediate values of h represent incomplete dominance. Weight these fledging probabilities by the probability that an individual has each genotype (i.e., the allele frequencies) to produce

$$\mathbf{L}_S = (a_t^S)^2 l_{tol} + 2(a_t^S)(1-a_t^S)(hl_{tol} + l_{wild} - hl_{wild}) + (1-a_t^S)^2 l_{wild}$$

Note that this is the inner product of g_t^S and l. $\mathbf{L}_A$ is defined the same way, substituting a_t^A in for a_t^S. This means that $\mathbf{L}_S = \langle g_t^S, l\rangle$ and $\mathbf{L}_A = \langle g_t^A, l\rangle$. Thus, our functions $\mathbf{L}_S$ and $\mathbf{L}_A$ are simply the inner products of genotypic frequencies and fledging probabilities. This means that our recruitment function is:

$$\begin{aligned} S_{t+1} &= \mathbf{b}_S S_t^b + \mathbf{b}_A A_t^b \\ &= \frac{\sigma_H \left(f_S S_t^b \langle g_t^S, l\rangle + f_A A_t^b \langle g_t^A, l\rangle \right)}{1 + \frac{1-d}{d\phi}(S_t + A_t)} \end{aligned} \tag{2.7}$$

where $\langle , \rangle$ denotes the inner product.

2.1.7 Change in allele frequency in subadults

The frequency of the tolerant allele is the number of copies of the allele divided by the total number of alleles in the population. Each tolerant homozygote has two copies of the allele, and each tolerant heterozygote has one copy of the allele. In fledgelings, the number of copies of the tolerant allele is

$$S_t^b f_S \big(2(a_t^S)^2 l_{tol} + 2a_t^S(1-a_t^S)(hl_{tol} + (1-h)l_{wild})\big)\sigma_H$$

Each fledgeling has two copies of this gene. Dividing the number of copies of the gene (i.e., twice the population size) in the subpopulation gives proportion of all subadult genes that have the tolerant allele, i.e., the tolerance allele frequency in subadults.

$$a_t^S = \frac{2S_t^b f_S \sigma_H \left((a_t^S)^2 l_{tol} + 2(a_t^S)(1-a_t^S)(hl_{tol} + (1-h)l_{wild})\right)}{2S_t^b f_S \sigma_H \left((a_t^S)^2 l_{tol} + 2a_t^S(1-a_t^S)(hl_{wild} + (1-h)l_{tol}) + (1-a_t^S)^2 l_{wild}\right)}$$
$$= \frac{(a_t^S)^2 l_{wild} + a_t^S(1-a_t^S)l_{wild}}{\mathbf{L}_S}$$

Define the following vectors

$$g_t'^S = \begin{bmatrix}(a_t^S)^2 & (a_t^S)(1-a_t^S) & 0\end{bmatrix} \qquad g_t'^A = \begin{bmatrix}(a_t^A)^2 & (a_t^A)(1-a_t^A) & 0\end{bmatrix}$$

the l_1 norms of these vectors are one half the number of copies of the tolerant alleles in the subadult and adult populations, respectively. We use one half of the number of copies of the allele because this allows us to divide by the total population size (rather than the number of copies of alleles, i.e., twice the population size) to get allele frequencies.

We use population sizes to calculate the expected number of copies of an allele from one generation to the next. Since selection has not yet acted when eggs are first laid, the frequency of the tolerant allele in eggs is the same as the frequency in breeding parents (see 2.1.5). Scaling the vector $g_t'^S$ by $f_S \sigma_H S_t^b$ (where f_S is fecundity of adults and σ_H is overwinter survival) gives the number of copies of the tolerant allele in eggs laid by second years, and likewise for the vector $f_A \sigma_H A_t^b g_t'^A$. $S_t^b \langle g_t'^S, l\rangle$ and $A_t^b \langle g_t'^A, l\rangle$ give the number of copies of the tolerant allele in fledged offspring by weighting fledging probabilities by genotypic frequencies.

To calculate the frequency of alleles in the subadult population in the next generation, divide the half-number of copies of the tolerant allele in offspring by the total number of individuals born. The half-number of alleles is

$$\frac{f_S \sigma_H S_t^b \langle g_t'^S, l\rangle}{1 + \frac{1-d}{d\phi}(S_t + A_t)} + \frac{f_A \sigma_H A_t^b \langle g_t'^A, l\rangle}{1 + \frac{1-d}{d\phi}(S_t + A_t)}$$
$$= \frac{\sigma_H \left(f_S S_t^b \langle g_t'^S, l\rangle + f_A A_t^b \langle g_t'^A, l\rangle\right)}{1 + \frac{1-d}{d\phi}(S_t + A_t)}$$

The total subadult population size in the next generation is,

$$S_{t+1} = \frac{\sigma_H \left(f_S S_t^b \langle g_t^S, l\rangle + f_A A_t^b \langle g_t^A, l\rangle\right)}{1 + \frac{1-d}{d\phi}(S_t + A_t)\big) S_{t+1}}$$

We note that this equation uses g_t^S and g_t^A instead of $g_t'^S$ and $g_t'^S$ used for the

half-number of alleles. The allele frequency in subadults in the next generation is:

$$\begin{aligned} a_{t+1}^S &= \frac{f_S S_t^b \langle g_t'^S, l\rangle + f_A A_t^b \langle g_t'^A, l\rangle}{f_S S_t^b \langle g_t^S, l\rangle + f_A A_t^b \langle g_t^A, l\rangle} \\ &= \frac{\sigma_H \left(f_S S_t^b \langle g_t'^S, l\rangle + f_S A_t^b \langle g_t'^A, l\rangle\right)}{\left(1 + \frac{1-d}{d\phi}(S_t + A_t)\right) S_{t+1}} \end{aligned} \tag{2.8}$$

These two expressions are equivalent; the bottom one is implemented in model code.

2.1.8 Change in allele frequency in adults

We assume that the probability of survival into adulthood is independent of genotype. The number of copies of the allele in adults in the next time step is $\sigma_S S_t a_t^S + \sigma_A A_t a_t^A$. Dividing this by the whole population size gives the frequency in adults in the next generation.

$$\begin{aligned} a_{t+1}^A &= \frac{\sigma_S S_t a_t^S + \sigma_A A_t a_t^A}{\sigma_S S_t + \sigma_A A_t} \\ &= \frac{\sigma_S S_t a_t^S + \sigma_A A_t a_t^A}{A_{t+1}} \end{aligned} \tag{2.9}$$

2.1.9 Carrying Capacity

A patch's breeding population can not exceed its number of nesting sites. We assume that there is no territory compacting, meaning that the number of nesting sites stays constant over time. This means that there is an upper limit to the number of breeders in a population, and thus an upper limit to population growth. In the absence of dispersal, if allele frequencies are held constant over time, equilibrium population size can be found analytically. Call the respective number of subadults and adults at equilibrium S^* and A^*. At this equilibrium, the following conditions are satisfied:

$$\begin{aligned} S^* &= \min\left(1, \frac{\phi}{S^* + A^*}\right) \frac{f_S \mathbf{L}_S^* \sigma_H S^* + f_A \mathbf{L}_A^* \sigma_H A^*}{1 + \frac{1-d}{d\phi}(S_t + A_t)} \\ A^* &= \sigma_S S^* + \sigma_A A^* \end{aligned}$$

$\mathbf{L}_S^*$ and $\mathbf{L}_A^*$ are fledging probabilities at equilibrium (and therefore constant).

In the absence of over dominance, mutation, [22], or dispersal, directional selection will move each subpopulation to fixation of its most advantageous allele.

Solving for A^* in terms of S^* gives:

$$A^* = \frac{\sigma_S S^*}{1-\sigma_A}$$

Note, now, that $A^* + S^* = S^* \left(1 + \frac{\sigma_S}{1-\sigma_A}\right)$. Designate $x = \frac{\sigma_S}{1-\sigma_A}$. and substitute this back into the equation for S.

$$S^* = \min\left(1, \frac{\phi}{S^*(1+x)}\right) \frac{\sigma_H S^* (f_S \mathbf{L}_S^* + f_A \mathbf{L}_A^* x)}{1 + \frac{1-d}{d\phi} S^*(1+x)}$$

Consider R', the net reproductive output of a bird in the absence of density-dependence; R' is a dimensionless parameter combining fecundity and survivorship; Caswell defines net reproductive output as $\sum_i^\infty \sigma_i b_i$ [8]. For our birds,

$$\begin{aligned} R' &= f_S \mathbf{L}_S^* \sigma_H + \sigma_S f_A \mathbf{L}_A^* \sigma_H + \sigma_S \sigma_A f_A \mathbf{L}_A^* \sigma_H + \sigma_S \sigma_A^2 f_A \mathbf{L}_A^* \sigma_H + \dots \\ &= \sigma_H \left(f_S \mathbf{L}_S^* + \sigma_S f_A \mathbf{L}_A^* \left(1 + \sigma_A + \sigma_A^2 + \dots\right)\right) \\ &= \sigma_H \left(f_S \mathbf{L}_S^* + \frac{\sigma_S f_A \mathbf{L}_A^*}{1-\sigma_A}\right) \\ &= \mathbf{b}_S + x\mathbf{b}_A \end{aligned}$$

Thus,

$$S^* = \min\left(1, \frac{\phi}{S^*(1+x)}\right) \frac{S^* R'}{1 + \frac{1-d}{d\phi} S^*(1+x)}$$

This is a piecewise function, depending on $\min\left(1, \frac{\phi}{S^*(1+x)}\right)$. This piecewise function can also be thought of as depending on R'. Consider the case when when there are fewer birds than there are breeding territories, $\frac{\phi}{S^*(1+x)} > 1$. Then, $\min\left(1, \frac{\phi}{S^*(1+x)}\right) = 1$, and

$$
\begin{aligned}
S^* &= \frac{S^* R'}{1 + \frac{1-d}{d\phi} S^*(1+x)} \\
R' &= \frac{1-d}{d\phi} S^*(1+x) \\
S^*(1+x) &= \frac{\phi(R'-1)}{\frac{1-d}{d}} \\
S^* &= \frac{d\phi(R'-1)}{(1-d)(1+x)}
\end{aligned}
$$

If population is larger than the number of breeding territories, $\min\left(1, \frac{\phi}{S^*(1+x)}\right) = \frac{\phi}{S^*(1+x)}$. Thus,

$$
\begin{aligned}
S^* &= \frac{\phi}{S^*(1+x)} \frac{S^* R'}{1 + \frac{1-d}{d\phi} S^*(1+x)} \\
S^* &= \frac{\phi}{1+x} \frac{R'}{1 + \frac{1-d}{d\phi} S^*(1+x)} \\
\frac{R'\phi}{1+x} &= S^* + (S^*)^2 \frac{(1-d)(1-x)}{d\phi}
\end{aligned}
$$

The quadratic formula tells us that:

$$
S^* = \frac{\phi}{2(1+x)\frac{1-d}{d}} \left(-1 + \sqrt{1 + 4R' \frac{1-d}{d}} \right)
$$

In both equations, as $\frac{\phi}{S^*(1+x)}$ limits to 1 (i.e., as population approaches carrying capacity from above or below), R' approaches $\frac{1-d}{d}$. Note that as $\frac{\phi}{S^*(1+x)}$ approaches 1, S^* approaches $\frac{\phi}{1+x}$; as it approaches $\frac{\phi}{1+x}$ from below (i.e., for populations smaller than the number of breeding territories)

$$
\begin{aligned}
\frac{\phi}{1+x} &= \frac{(R'-1)\phi}{\frac{1-d}{d}(1+x)} \\
1 &= \frac{(R'-1)}{\frac{1-d}{d}} \\
R' &= 1 + \frac{1-d}{d} = \frac{1}{d}
\end{aligned}
$$

Limiting from above, i.e., for populations larger than carrying capacity,

$$\begin{aligned}
\frac{\phi}{1+x} &= \frac{\phi}{2(1+x)\frac{1-d}{d}}\left(-1+\sqrt{1+4R'\frac{1-d}{d}}\right) \\
\sqrt{1+4R'\frac{1-d}{d}} &= \frac{2}{d}-1 \\
4R'\frac{1-d}{d} &= \frac{4-4d}{d^2}-\frac{4}{d} \\
R'\frac{1-d}{d} &= \frac{1-d}{d}\frac{1}{d} \\
R' &= \frac{1}{d}
\end{aligned}$$

This means that S^* and A^* are actually continuous functions of the dimensionless parameter R'. If $R' > \frac{1}{d}$, then the equilibrium population will have floaters, and if $R' < \frac{1}{d}$ the equilibrium population will have no floaters. This threshold for reproductive rate is intuitive; as the total population size approaches the number of breeding territories (ϕ), net reproductive rate approaches dR' (where R' is the reproductive rate in the absence of density dependence, i.e., as the population size approaches 0). In order to have a population exceed the total number of breeding territories, $dR' > 1$, or $R' > \frac{1}{d}$. Thus, the threshold for reproductive rate needed to guarantee floaters increases with the inverse of d, which is a measure of density dependence. Indeed, numerical simulations of populations with zero dispersal show that for subpopulations with life history parameters such that $R' > \frac{1}{d}$, equilibrium population sizes exceed the number of breeding sites in each patch, subpopulations with $\frac{1}{d} > R' > 1$ leads to equilibrium populations smaller than the number of breeding sites in each patch, and $R' < 1$ leads to a population crash.

At steady state,

$$S = \begin{cases} 0 & \text{if } 0 < R' < 1 \\ \frac{d\phi(R'-1)}{(1-d)(1+x)} & \text{if } 1 < R' \leq \frac{1}{d} \\ \frac{\phi}{2(1+x)\frac{1-d}{d}}\left(-1+\sqrt{1+4R'\frac{1-d}{d}}\right) & \end{cases} \tag{2.10}$$

Additionally, at steady state $A = \frac{\sigma_S}{1-\sigma_A}S$.

2.2 Spatial Structure

Now we move from considering a single population in isolation to a fragmented population made up of n discrete patches. In the case of the Shenandoah Val-

ley, some of these subpopulations are along the South River, and therefore are contaminated with mercury [10], while other subpopulations are further from the river where presumably mercury exists in much lower dosages or is not physically present.

2.2.1 Floater Dispersal

We assume that birds disperse between habitats in search of open nests. It follows that only floaters will disperse, since birds in the breeding population already have nests. Floaters, then, play a key role in this model by allowing gene flow. The model records bird abundances in each patch at the end of each time step; section 2.1.2 describes how the model tracks breeding and non-breeding birds. Thus, in each patch at each time step, the model has a value representing abundance of non-breeding "floater" birds. However, if allowing for dispersal, these floaters can move between patches in search of open nests; thus, a floater can begin a time step in one patch, then move to another patch in order to compete for nests there. In this model, birds may only disperse once per time step.

Describing dispersal requires describing where these floaters are likely to move. This model assumes that the probability of an individual dispersing from a patch i to a patch j is dependent on two things: the distance between i and j, and the number of open breeding territories in j (this means dispersal follows a "gravity model"). The first assumption is quite naive when it comes to bird dispersal models, as landscape features tend to affect dispersal behavior, and that response to such features varies among birds [19]. This problem can be solved by using *effective* distances between patches, which incorporate landscape features that modify dispersal into parameters that determine dispersal likelihood. While this makes inter-patch distances difficult to calculate in empirically in extant natural systems, it is easy to incorporate barriers to dispersal into theoretical studies or hypothetical simulations. We also use the simplifying assumption that dispersal depends, at least in part, on the number of open breeding territories. Our other assumption is that birds disperse towards sites that have more open breeding sites. This assumes that perhaps bird are equally likely to visit each patch, but are most likely to stop and settle if it sees an open patch; the probability that it sees an open patch is proportional to the number of open patches.

Let d_{ij} be a measure of effective distance from patch j to patch i, with $d_{jj} = 0$ for all j (that is, there is no travel cost to stay within a patch). We will make the simplifying assumption in our analysis that $d_{ij} = d_{ji}$ for every pair i, j, although it is not a necessary condition for the model. This model tracks the size of a whole population of birds by subdividing the population into discrete subpopulations

(patches) and keeping track of the number of birds in each patch. In order to simulate a bird competing for multiple nests in multiple patches at once, we proportion the floater population of each patch based on the probability that a bird will try to compete in each patch, then add these values to populations in their destination patches. This can be done by matrix multiplication and is explained in equation 2.11.

The pre-dispersal number of floater adults in patch i at time t is $A^f_{i,t} = A_{i,t} - A^n_{i,t}$. Assume that the distance a bird is willing to fly to compete for a nest decays exponentially; this is featured in several other models of dispersal [1]. Let $e^{-\alpha d_{ij}}$ be a decay rate of a bird's willingness to move to compete in patch i from patch j, where α is a nonnegative parameter. Higher α indicates lower chance of competing over long distances; thus α can be thought of as a quantification of barriers to dispersal. Furthermore, assume that the likelihood a bird will invade a territory is directly proportional to the number of open spots in the destination territory; that is, a bird in patch j will be more likely to compete in territory i the higher ϕ'_i is. The probability that a bird in j will compete in i, then, is $\frac{\phi'_i e^{-\alpha d_{ij}}}{\sum_{k\in\mathcal{N}_j} \phi'_k e^{-\alpha d_{kj}}}$, where $\mathcal{N}_j$ is the set of breeding territories reachable from j. The expected number of floaters resident in patch j who will disperse into i would be $A^f_{i,t-1}$ (the number of birds at the end of time step $t-1$ is the same as the number at the beginning of time step t) times the term above. The total number of floaters in patch i after dispersal, which we will denote by $\hat{A^f_i}$, is the sum:

$$\hat{A}^f_{i,t} = \sum_{j\in\mathcal{N}_j} \frac{\phi'_i e^{-\alpha d_{ij}}}{\sum_{k\in\mathcal{N}_j}(e^{-\alpha d_{kj}}\phi'_k)} A^f_{j,t-1} \tag{2.11}$$

If A^f_t is an $n \times 1$ vector where the ith entry represents the number of floater adults before dispersal in subpopulation i at time t, and $\hat{\mathsf{A}}^f_t$ is likewise an $n \times 1$ vector of post-dispersal adult floater population sizes, then the above can be rewritten as:

$$\hat{\mathsf{A}}^f_t = \mathsf{M}\mathsf{A}^f_{t-1}$$

where M is an $n \times n$ movement matrix such that:

$$m_{ij} = \frac{\phi'_i e^{-\alpha d_{ij}}}{\sum_{k\in\mathcal{N}_j}(e^{-\alpha d_{kj}}\phi'_k)} \tag{2.12}$$

Note that each column in M sums to 1. Also note that M changes over time, as

it is dependent on ϕ' which also changes over time. The value in equation 2.11 is used for $A^f_{i,t}$ in equations 2.5 and 2.6 to determine the number of breeding birds in each patch. A proportion of these floaters will gain nests in i; the remaining floaters will remain floaters in i until the next time step. The number of floaters in patch i after nest settlement, then, is:

$$A^f_{i,t} = \hat{A}^f_{i,t} - \min\left(1, \frac{\phi'_i}{\hat{A}^f_{i,t} + \hat{S}_{i,t}}\right)\hat{A}^f_{i,t} \tag{2.13}$$

Subadults will follow the same rules as adults for dispersal. However, since subadults have never competed for nests before, all subadults can disperse. Thus, the post-dispersal number of subadults can be found using a the matrix equation:

$$\hat{\mathsf{S}}_t = \mathsf{M}\mathsf{S}_{t-1}$$

where M is the same as defined above for adults. These values $\hat{A}^b_t$ and $\hat{S}^b_t$ are used for the recruitment function in equation 2.7 and for calculating allele frequencies in equations 2.8 and equations 2.9.

2.2.2 Allele Frequencies

Consider a hypothetical subpopulation i where floaters and breeders both have an allele frequency a_i. If a proportion of floaters emigrate and floaters from other subpopulations immigrate into i, the breeding population frequency will still be a_i (and the subadults in the next generation will have a frequency as a function of a_i) while the floating population will have an allele frequency that is a function of the frequencies of immigrants. The allele frequencies of floaters and breeders will inevitably diverge, especially in situations where floaters rarely overtake territories from established breeders (ι is close to 1) and the two groups do not mix [13]. This necessitates keeping track of the allele frequencies of breeders and nesters separately.

Define $a^A_{f,i,t}$ to be the allele frequency of adult floaters in subpopulation i at time t before dispersal, $a^A_{b,i,t}$ to be the allele frequency of adult breeders in subpopulation i at time t before dispersal, and $a^S_{i,t}$ to be the allele frequency of subadults in subpopulation i at time t (remember that before nest settlement there are no subadult breeders; for this reason we do not distinguish between subadults floaters and breeders). These variables with a hat (e.g., $\hat{a}^A_{f,i,t}$) represent the allele frequencies after dispersal.

Because nest resettlement involves mixing floater subpopulations, the allele frequency in the breeding population of i after resettlement is a function of the

breeding population frequency in i before resettlement and the floater population frequencies in all other patches. The expected number of copies of the tolerant allele, as before, is the half-number of copies of the tolerant allele divided by the total population. Before resettlement of nests, the allele frequencies are the same as at the end of the $t-1$ step. In subadults, the half-number of alleles is

$$\sum_{j\in\mathcal{N}_j} \frac{\phi_i' e^{-\alpha d_{ij}}}{\sum_{k\in\mathcal{N}_j}\phi_k' e^{-\alpha d_{kj}}} S_{j,t-1} a^S_{j,t-1}$$

the frequency of alleles in subadults after dispersal in patch i is

$$\hat{a}^S_{i,t} = \frac{1}{\hat{S}_{i,t}} \sum_{j\in\mathcal{N}_j} \frac{\phi_i' e^{-\alpha d_{ij}}}{\sum_{k\in\mathcal{N}_j}\phi_k' e^{-\alpha d_{kj}}} S_{j,t-1} a^S_{j,t-1} \tag{2.14}$$

where $\hat{S}_{i,t}$ is the number of subadults after dispersal. Using a^S_t to denote an $n\times 1$ vector of subadult allele frequencies before dispersal at time t, the above can be rewritten as

$$\begin{aligned}\hat{\mathsf{a}}^S_{i,t} &= \frac{\mathsf{M}_{(i,)}(\mathsf{S}_{t-1}\circ \mathsf{a}^S_{t-1})}{\hat{\mathsf{S}}_t} \\ &= \frac{\mathsf{M}_{(i,)}(\mathsf{S}_{t-1}\circ \mathsf{a}^S_{t-1})}{\mathsf{M}_{(i,)}\mathsf{S}_{t-1}}\end{aligned} \tag{2.15}$$

where $\mathsf{M}_{(i,)}$ is the ith row of the dispersal matrix M and $\circ$ denotes the Hadamard product. Equation 2.15 is implemented in code. A similar equation can be derived for the adult floater allele frequency. Since the half-number of alleles in adult floaters is

$$\sum_{j\in\mathcal{N}_j} \frac{\phi_i' e^{-\alpha d_{ij}}}{\sum_{k\in\mathcal{N}_j}\phi_k' e^{-\alpha d_{kj}}} A^f_{j,t-1} a^A_{f,j,t-1}$$

the frequency of alleles of floaters competing in patch i at time t is:

$$\hat{a}^A_{f,i,t} = \frac{1}{\hat{A}^f_{i,t}} \sum_{j\in\mathcal{N}_j} \frac{\phi_i' e^{-\alpha d_{ij}}}{\sum_{k\in\mathcal{N}_j}\phi_k' e^{-\alpha d_{kj}}} A^f_{j,t-1} a^A_{f,j,t-1} \tag{2.16}$$

where $\hat{A}^f_{i,t}$ is the same as defined in equation 2.11. Using matrix multiplication, we can rewrite this as:

$$\begin{aligned}\hat{\mathsf{a}}^A_{i,t} &= \frac{\mathsf{M}_{(i,)}(\mathsf{A}^f_{t-1} \circ \mathsf{a}^A_{t-1})}{\hat{\mathsf{A}}_t} \\ &= \frac{\mathsf{M}_{(i,)}(\mathsf{A}^f_{t-1} \circ \mathsf{a}^A_{t-1})}{\mathsf{M}_{(i,)}\mathsf{A}_{t-1}}\end{aligned} \tag{2.17}$$

Equation 2.17 is implemented in code.

When nests are settled by adult floaters and subadults (see 2.1.3), the settling birds (breeders) are a randomly selected proportion of the birds competing for these nests. The pool of breeding adults is made up of some birds who nested in the previous time interval ($A^n = \iota A^b_{i,t-1}$) and floater adults ($\hat{A}^f_{i,t}$). The frequency of alleles in breeding adults, then, is:

$$\hat{a}^A_{b,i,t} = \frac{A^n a^A_{b,i,t-1} + \min\left(1, \frac{\phi'_i}{\hat{S}_{i,t}+\hat{A}^f_{i,t}}\right)\hat{A}^f_{i,t}\hat{a}^A_{f,i,t}}{A^n + \min\left(1, \frac{\phi'_i}{\hat{S}_{i,t}+\hat{A}^f_{i,t}}\right)\hat{A}^f_{i,t}} \tag{2.18}$$

We assume that when floaters gain nests during resettlement, they do so at random. Since there is no difference among genotypes in ability to gain nests, we may assume that the allelic frequency of floaters gaining nests in i are the same as the frequencies of the entire floater subpopulation in i, as the floaters who gain nest can be considered a random sample of all floaters. This is why we are allowed to use the term $\hat{a}^A_{f,i,t}$ for describing allelic frequencies of breeders-turned-nesters in equation 2.18. These values $\hat{a}^S_{f,i,t}, \hat{a}^S_{b,i,t}, \hat{a}^A_{f,i,t}, \hat{a}^A_{b,i,t}$ are used to calculate fledging probabilities $\mathbf{L}_S$ and $\mathbf{L}_A$ for the allele frequencies of offspring as described in section 2.1.6, as well as for calculating change in allele frequencies due to survivorship in section 2.1.8.

2.3 Algorithm

The script begins by initializing variables and a dispersal matrix.

function SETUP(Number contaminated patches, number uncontaminated patches, carrying capacity list, distance matrix)

Create global list of patch sizes, ϕ, distance matrix, variables A, S
S^b_0, S^f_0, A^b_0, $A^f_0 \leftarrow$ values from equation 2.10
$a^S_{b,0}, a^S_{f,0}, a^A_{b,0}, a^A_{f,0} \leftarrow \frac{2}{A^f_0}$ for all contaminated patches
$a^S_{b,0}, a^S_{f,0}, a^A_{b,0}, a^A_{f,0} \leftarrow 0$ for all uncontaminated patches
$t = 1$

end function

Next, the script loops over three procedures, each meant to simulate a different part of a bird's annual life cycle.

repeat

function RESETTLE(time step t)

Locally store number of nesters A^n, open spots ϕ', dispersal matrix M

Calculate $\hat{S}_t^b, \hat{S}_t^f, \hat{A}_t^b, \hat{A}_t^f$ using equation 2.13

Update $\hat{a}_{b,t}^S, \hat{a}_{f,t}^S, \hat{a}_{b,t}^A, \hat{a}_{f,t}^A$ using equations 2.16 and 2.18

Calculate ϵ from equation 3.1

end function

function BIRTH(time step t)

Calculate S_t from equation 2.7

Update $a_{b,t}^S$, $a_{f,t}^S$ using equation 2.8

end function

function MATURE(time step t)

Update $a_{b,t}^A$, $a_{f,t}^A$ from equation 2.9

Update A_t^b, A_t^f from equation 2.2

end function

Calculate weighted averages of allele frequencies in each patch

time $\leftarrow t + 1$

until All allele frequency change smaller than 10^{-5} and all population changes smaller than 10^{-5}

Record t^*, time until equilibrium, and $S_{i.t^*}, A_{i,t^*}, a_{i,t^*}^S, a_{i,t^*}^A, R'$

The loop terminates at a state close to equilibrium (exact equilibrium is near impossible to reach due to computational precision). At this point, several output variables are stored. Code is implemented using R 3.0.1.

Chapter 3

Model Analysis

Although the model is equipped to handle an arbitrary number of patches, to gain an initial understanding we will only do analysis on two-patch landscapes. Each landscape will have one contaminated patch and one uncontaminated patch; this removes landscape configuration from consideration as an input variable.

3.1 Global Sensitivity Analysis

In searching for conditions for migration load and local maladaptation, we are ultimately asking questions about model sensitivity, in other words, what has the greatest effects on various model outputs, specifically allele frequencies? We expect that our model will produce variation in allele frequencies and population sizes, and global sensitivity analysis provides a framework for determining which input variables cause the greatest variation in model output [32]. Partial rank correlation (PRC) quantifies the effects of individual input parameters on individual model outputs after removing the effects and interactions of all other variables [20] (whereas partial correlation describes linear relationships between variables and outputs, partial rank correlation describes this relationship on rank transformed data) [24]. The output of PRC is a partial rank correlation coefficient, a statistic on the range $[-1, 1]$, where positive numbers indicate a positive relationship between a parameter and response variable, negative numbers indicate a negative relationship, and higher magnitudes suggest a stronger relationship. The statistical significance of PRC coefficients can be assessed using student's t-test [24].

We isolate eleven parameters of interest to use in PRC. For each parameter, we choose three values. These are shown in Table 3.1. Median life history parameter values $f_S, f_A, \sigma_S, \sigma_A$, as well as the median value for d were recommended by Cristol (personal communication); for each parameter value, a point above and below the values recommended were arbitrarily selected to allow study over a range of

parameter space. The values for parameters $l_{wild,c}$ and $l_{tol,u}$ come from laboratory experiments conducted by Varian-Ramos et. al., 2013 and 2014, on zebra finches. Cristol (personal communication) recommended .8 as a baseline fledging probability, and laboratory experiments found birds exposed to mercury suffered between a 20% and 50% reduction in fitness; 20% reductions coming from dosage concentrations observed in the Shenandoah, and 50% reductions coming from "worst case scenario" dosage concentrations. The value $\iota = 1$ corresponds to all nesting birds retaining their nests, and $\iota = 0$ corresponds to a "free for all" for all nests in each year. The value $h = 1$ corresponds to a dominant tolerant allele, $h = 0$ corresponds to a recessive tolerant allele, and $h = .5$ corresponds to incomplete dominance. For the value $\alpha = 0$, birds disperse between patches regardless of inter-patch distance (but still favor patches with more open breeding territories); $\alpha = 1$ and $\alpha = 2$ correspond to birds that increasingly favor nearby patches over distant patches, i.e., higher α means smaller dispersal radius.

The size of the contaminated patch, ϕ_c, is held constant at 10. With changes in ϕ_u, the uncontaminated patch can either be half as large, the same size as, or twice as large as the contaminated patch. The parameter σ_H is held constant at 0.3. In our two patch model, we let inter-patch distances $\delta_{cu} = \delta_{uc} = 1$. We bundled the effects of α, decay rate of dispersal likelihood, and inter-patch distances into a single parameter. However, in a hypothetical case with more patches, inter-patch distances (landscape configuration) and α may have different effects, and would thus have to be studied separately.

Testing all combinations of parameters would require 3^{11} trials of the model. To limit our sample size, we use the Latin Hypercube Sampling technique to generate 1500 distinct combinations [32]. These samples were created by creating a 1500×11 array, with a row for each parameter combination and a column for each parameter. Each column was filled with 500 copies of each of the three values chosen for each parameter. Each of these columns was then shuffled. This method treats all parameter combinations as equally likely and samples parameter combinations with replacement. Although biologically it is unrealistic to treat all parameter combinations as equally likely, we are using this technique simply to gain a complete understanding of the model, so that we may understand the relative effects of each parameter in the model. Thus we want to adequately sample the sample space, giving no preference to certain regions that may be more biologically relevant.

The model was run for each of these 1500 samples until allele frequencies and population sizes were near constant in both patches (change in allele frequency and change in population size were $< 10^{-5}$ for each subpopulation and stage class). Let the time equilibrium is reached be t^*. At time t^*, the model was stopped, and the following response variables were recorded for each patch:

Parameter	Parameter Meaning	Values		
f_S	Fecundity of subadults	.6	.9	1.2
f_A	Fecundity of adults	.9	1.2	1.5
σ_S	Survival probability of subadults	.2	.4	.6
σ_A	Survival probability of adults	.3	.5	.7
α	Rate of decay of dispersal probability with distance	0	1	2
ι	Nest guarantee parameter	0	.5	1
$l_{wild,c}$	Fledging probability of wild type bird exposed to mercury	.4	.52	.64
$l_{top,u}$	Fledging probability of tolerant bird in absence of mercury	.4	.52	.64
ϕ_u	Number of breeding territories in uncontaminated path	5	10	20
h	Heterozygosity modulator	0	.5	1
d	Proportion of reproductive output when all nests full	.84	.9	.96

Table 3.1: Eleven parameters chosen, with three values each, for latin hypercube sampling.

1. Population size relative to the equilibrium population size as defined in equation 2.10, i.e., $\frac{S_{t^*}+A_{t^*}}{S^*+A^*}$

2. Allele frequencies of subadults and adults weighted by their proportion of the total population, i.e., $\frac{S_{t^*}a^S_{t^*}+A_{t^*}a^A_{t^*}}{S_{t^*}+A_{t^*}}$

3. $R' = \sum_i \mathbf{b}_i\sigma_i$, our measure of net reproductive output in the absence of density dependence defined in 2.1.9. This is recorded for each patch, as well as a population-wide R' that is weighted by the size of each subpopulation. This R' value can also be interpreted as a measure of population fitness [16].

4. A measure of gene flow into each patch, given by

$$\epsilon_i = \frac{\sum_{j\neq i}\left(m_{ij}(S_j + A^f_j)\right)\min\left(1, \frac{\phi_i}{S_i+A^f_i}\right)}{\sum_j\left(m_{ij}(S_j + A^f_j)\right)\min\left(1, \frac{\phi_i}{S_i+A^f_i}\right) + A^n_i} \tag{3.1}$$

   The numerator of 3.1 is the proportion of dispersers into patch i (that did not originate in i) attaining nests; the denominator is the proportion of all dispersers into i attaining nests, plus the number of nesters. The numerator, then, is the number of breeders in i originating outside of i in that time step, and the denominator is the total number of breeders in i in that time step. This makes ϵ_i the proportion of breeders in i originating outside of i, which serves as an appropriate measure of gene flow.

3.2 Methods

We will perform PRC for each pair of parameter and response variable. This was done using the R package "sensitivity" [30]. Of our sample of trial runs, we perform PRC using the "pcc" command, calculating 95% confidence intervals via bootstrapping with $n = 100$ bootstrapped samples. We will group statistically significant ($p < .05$) data points with overlapping confidence intervals to assess the primary parameters that each of our output variables is most sensitive to.

Our first prediction posed in section 1.2 is that local selection pressure against the maladapted allele will decrease the frequency of that allele. To test this prediction, we varied the strength of selection pressure in each patch. Our first prediction will be tested by assessing PRC coefficients of selection pressures in a patch and allele frequencies at equilibrium in each patch. Model input in this case is l_{wild} in the contaminated patch and l_{tol} in the uncontaminated patch; thus, higher input variables correspond to weaker selection pressures (e.g greater l_{wild} in the contaminated patch corresponds to higher fledging probability of eggs of wild type parents exposed to mercury). Our prediction will be supported, then, if l_{wild} has a negative PRC coefficient, as higher fledging probability of wild type birds in a contaminated patch will lead to more wild type birds and thus fewer tolerant birds in that area, and that l_{tol} will have a positive PRC coefficient, as higher fledging probability of tolerant birds in an uncontaminated patch would lead to more tolerant birds in that patch. We will also perform PRC between these parameters and R', the net reproductive rate in each patch. We predict that the relationship between selection pressure and R' will be similar to the relationship between selection pressure and allele frequencies.

Our second prediction is that more gene flow will lead to higher prevalence of the locally maladaptive allele. We have defined a parameter ϵ_i for the amount of gene flow into patch i. We will perform Pearson product-moment correlation between ϵ and allele frequency in each patch. Our prediction will be supported of ϵ_c has a negative correlation with the frequency of the tolerance allele in the contaminated patch, as more gene flow into the contaminated patch means more wild type wild type birds breeding into the local population. For the same reason, we expect that ϵ_u will correlate positively with frequency of the tolerant allele in the uncontaminated patch. We will also study Pearson correlation between gene flow and R', where we predict a similar relationship to gene flow and allele frequency.

Our third prediction, that increasing parameters that encourage movement between patches will lead to more maladaptation, can be tested by assessing PRC coefficients of α (dispersal likelihood decay rate) and ι (which controls the number of open nests) with allele frequencies in each patch. Increasing α lowers the

probability of dispersal between distant patches. We expect, then, a positive PRC between α and the tolerance allele frequency in the contaminated patch and a negative PRC between α and the tolerance allele frequency in the uncontaminated patch. Increasing ι decreases the number of open nests to compete over in a time step. We predict a positive relationship between ι and the tolerant allele frequency in the contaminated patch, and a negative relationship between ι and the tolerant allele frequency in the uncontaminated patch. To relate this to fitness, we will also evaluate the PRC coefficients of α and ι with R'.

Our fourth prediction is that increasing fecundity and survival probabilities and decreasing the strength of density dependence will lead to more maladaptation by creating more floaters and thus more gene flow. First, to ensure that a relationship between these parameters and gene flow exist, we will examine PRC coefficients of $f_S, f_A, \sigma_S, \sigma_A$, and d with ϵ. If a relationship between these parameters and gene flow exist, we will assess PRC coefficients of these parameters with allele frequencies in each patch. We expect negative relationships between fecundity and survivorship parameters and tolerance frequency in the contaminated patch and positive relationships between these parameters and tolerance frequency in the uncontaminated patch. We also predict the the same will be true of the relationship between d and allele frequencies, as increasing d decreases density dependence and thus creates more floaters. Because R' is a function of these life history trait parameters, we expect that R' will be very strongly positively affected by them, but we will not count these relationships as significant. However, we will examine PRC coefficients of d with R' in each patch.

Our final prediction is that asymmetrical dispersal will lead to more local maladaptation. First, to see if there is a relationship between asymmetrical dispersal and fitness, we will study the Pearson correlation between each net gene flow into the contaminated patch, $\epsilon_c - \epsilon_u$, and allele frequencies and R'. If a relationship exists, we will assess PRC coefficients of ϕ_u on allele frequency and fitness in both patches, with the assessment that lower ϕ_u means that breeding grounds that favor wild type birds are rare relative to grounds that favor tolerant birds. Increasing ϕ_u should lead to more wild type birds in the whole population. We predict then that ϕ_u will have a negative PRC coefficient with the tolerant allele in both patches.

3.3 Initial Conditions

We want to simulate sudden appearance of a contaminant in a landscape full of wild-type birds. Before the model begins, all patches are assumed to be free of mercury. In the absence of mercury, directional selection will guide each patch to analytic steady state determined by equation 2.10 and fixation of the wild-type

allele.

Before the first time step of the model, the contaminant is introduced to select patches. Also, two mutant copies of the tolerant allele appear in each contaminated patch (uncontaminated patches remain at fixation of the wild type allele). The effects of this dose are constant over time; that is, the amount of contaminant does not increase or decrease. Contaminant does not move between patches; this means that patches that are uncontaminated remain uncontaminated throughout the whole experiment. We also assume a static landscape, where patches do not change in size and inter-patch distances do not change.

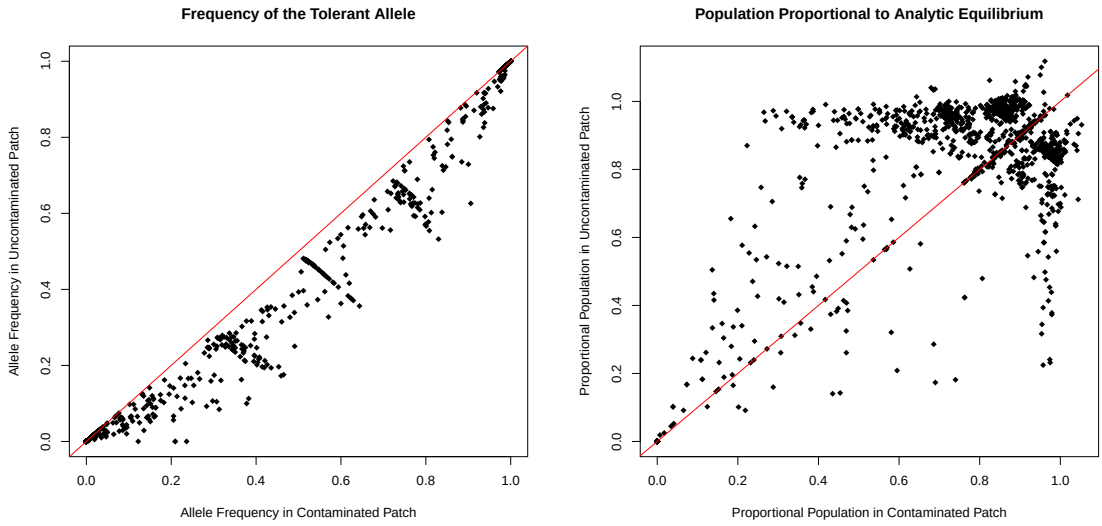
For parameter combinations where $R' < 1$ before the introduction of mercury, we do not run the model, as this corresponds to a population that will crash even without the introduction of mercury. Because of this, even though we use 1500 parameter combinations with Latin Hypercube Sampling, we may have fewer than 1500 data points.

3.4 Results

Because certain parameter combinations yielded $R' < 1$, 1197 data points were generated by running the simulation until steady state.

Steady state frequencies of the tolerant allele in each patch are plotted in figure 3.1a. Pearson's ρ gives a correlation between frequencies in patches of .991 with $p < 2.2 * 10^{-16}$, 1195 degrees of freedom. Maximum difference between allele frequencies in all simulations was .298. Median tolerant allele frequency was .048 in the contaminated patch and .029 in the uncontaminated patch. 250 parameter combinations out of 1197 (20.9%) had tolerant allele frequency above .95 in the contaminated patch. 243 (20.3%) had tolerant allele frequency above .95 in the uncontaminated patch; all of these 243 data points with high tolerance in the uncontaminated patch are included in the 250 points with high tolerance in the contaminated patch, suggesting that it is highly unlikely that a tolerance is prevalent in an uncontaminated area but not a contaminated one.

Median population size was 83.9% of equilibrium size in the contaminated patch and 90.3% of equilibrium size in the uncontaminated patch. Pearson's ρ for population sizes in the two patches was .725, $p < 2.2 * 10^{-16}$. In 63 trials (5%), population in both patches was less than 1% of the equilibrium population size in each patch; there was one additional trial where only the contaminated patch's population was below 1% of its equilibrium level and two cases of the converse. These trials can be thought of as population crashes. The weighted fitness parameter R' of the whole population was below 1 for 55 of these 63 instances of population crashing. The maximum whole-population R' value for instances of



(a) Tolerance allele frequency. $\rho = .991$

(b) Relative population size. $\rho = .725$

Figure 3.1: Scatter plots allele frequencies (a) and population relative to analytic steady state (b) for 1197 model trials. Contaminated patch on x axis, uncontaminated patch on y axis. Red line corresponds to the line $y = x$

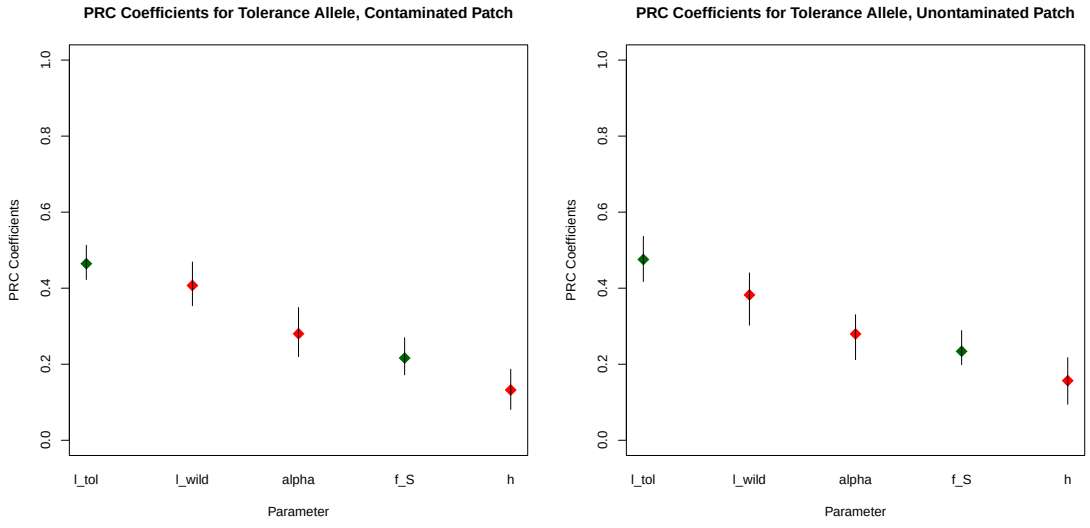
population crash was 1.074.

PRC coefficients for each of combination of parameter and response variable are listed in Table 3.2. For the tolerance frequency in both patches, the significant variables were (in order of decreasing magnitude) fledging probability of tolerant birds in the uncontaminated patch, fledging probability of wild type birds in the contaminated patch, dispersal decay rate, fecundity of subadults, and heterozygosity. These are shown in figures 3.2a and 3.2b.

Population sizes, which are divided by the analytically derived equilibrium population size in equation 2.10 to control for differences in patch size and life history parameters, show significant rank correlation with a number of parameters. Population size in the contaminated patch had a significant relationship with all parameters but density dependence, and population size in the uncontaminated patch had a significant relationship with all parameters but invasion probability, size of the uncontaminated patch, and density dependence. The strongest correlation in the contaminated patch was with the fledging probability of tolerant birds in the uncontaminated patch; the strongest correlation in the uncontaminated patch was with the fledging probability of wild-type birds in the contaminated patch. Life history traits, dispersal decay rate, and local maladaptive fledging probability had the next strongest correlations. Dispersal decay rate had a negative relationship in

	Pop_c	Pop_u	Tol_c	Tol_u	R_c	R_u	R_{total}	ϵ_c	ϵ_u	$\epsilon_c - \epsilon_u$
f_s	.277*	.111*	.216*	.234*	.606*	.610*	.758*	-.231*	.155*	-.013
f_a	.206*	.324*	.043	.051	.502*	.624*	.696*	-.053*	.011	-.029
σ_s	.256*	.394*	.032	.027	.716*	.846*	.891*	-.314*	-.229*	-.034
σ_a	.147*	.353*	-.058	-.038	.641*	.843*	.861*	-.272*	-.280*	-.007
α	-.210*	.346*	-.280*	-.280*	-.268*	.199*	-.063*	-.865*	-.908*	.029
ι	.175*	-.046	.035	-.002	.117*	-.150*	-.065*	-.657*	-.681*	.037
$l_{wild,c}$	.216*	.631*	-.407*	-.382*	.141*	.399*	.337*	-.076*	.008	-.042
$l_{tol,u}$	.444*	-.180*	.464*	.476*	.420*	-.284*	.241*	-.204*	.199*	.035
ϕ_u	-.069*	-.026	-.036	-.034	-.061	.009	-.066*	.058	-.004	-.018
h	.110*	-.109*	-.132*	-.157*	.167*	-.087*	.100*	.149*	.114*	.002
d	.010	.001	.022	.025	-.008	-.050	.029	-.000	.019	-.128*

Table 3.2: Mean bootstrapped PRCC values ($n = 100$) for parameter inputs (rows) on response variables (columns). $*$ is significance to the .05 level.



(a) Contaminated patch

(b) Uncontaminated patch

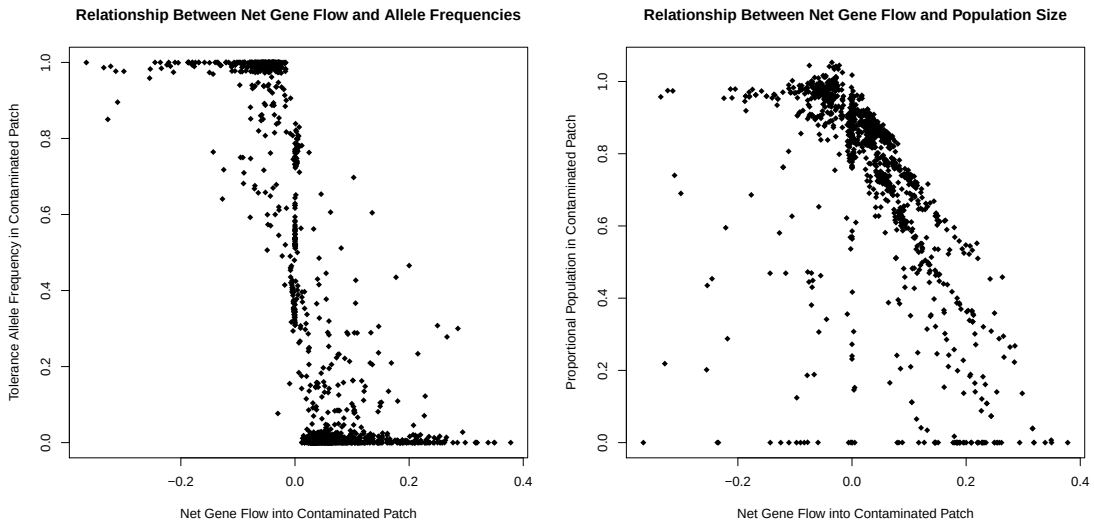
Figure 3.2: Significant ($p < .05$) PRC coefficients and confidence intervals for tolerant allele in each patch. Height corresponds to magnitude of coefficient, color corresponds to effect (green positive, red negative).

the contaminated patch and a positive relationship in the contaminated patch.

Reproductive rates R_c, R_u and R_{total} were most sensitive to life history parameters. These parameters were also sensitive to maladaptive fledging probabilities and dispersal decay rate; in both patches, the correlation with fledging probability in the opposite patch was stronger than correlation with local fledging probability. Also, just as with population sizes, dispersal has a negative effect on reproductive rate in the contaminated patch and a positive effect on reproductive rate in the uncontaminated patch. R_{total} was positively correlated with fledging probability in both patches.

Gene flow into the contaminated patch, ϵ_c was sensitive to all parameters except for size of the uncontaminated patch and density dependence. Gene flow into the uncontaminated patch, ϵ_u was sensitive to all parameters except for size of the contaminated patch, density dependence, wild type fledging probability in the contaminated patch, and fecundity of adults. For ϵ_c, all significant parameters had negative correlations. For ϵ_u, tolerant fledging probability in the uncontaminated patch, fecundity of subadults and heterozygosity had weak positive correlations, and all other parameters had stronger negative correlations. Net gene flow into the contaminated patch was only significantly correlated with strength of density dependence, a weak negative relationship.

Gene flow into the contaminated patch, ϵ_c, had a Pearson correlation coefficient (ρ) of $-.088$ with allele frequency in the contaminated patch ($p < .005$). Likewise, gene flow into the uncontaminated patch ϵ_u, had Pearson correlation coefficient of .384 with allele frequency in the uncontaminated patch ($p < 2.2 * 10^{-16}$). ϵ_c, when compared with reproductive output in the contaminated patch, yielded $\rho = -.181$, $p < .001$. ϵ_u when compared with reproductive output in the uncontaminated patch yielded $\rho = -.206$, $p < .001$. Net gene flow into the contaminated patch had a Pearson correlation coefficient $\rho = -.738$ with tolerant allele frequency in the contaminated patch ($p < 2.2 * 10^{-16}$), $\rho = -.517$ with fitness in the contaminated patch R_c ($p < .001$) and $\rho = -.225$ with the total weighted fitness R_{total} ($p < 2.2*10^{-16}$). Correlation between net gene flow and allele frequencies and population sizes are shown in figures 3.3a and 3.3b.



(a) Net gene flow and frequency tolerance. (b) Net gene flow and population size

Figure 3.3: Relationship between net gene flow into the contaminated patch, $\epsilon_c - \epsilon_u$, and allele frequencies (a) and population sizes (b).

Chapter 4

Conclusion

4.1 Discussion

We built a model for the spread of a mercury tolerance gene in a spatially structured bird population. In order to assess likelihood of tolerance fixation and migration load in a two-patch scenario, we selected model parameters assumed to correlate with allele frequency, population size and reproductive rate, and performed global sensitivity analysis on 1197 simulated outcomes.

Within the domain parameter space we studied, we found that half of trials reached a steady state with the tolerance allele at less than 5% in both patches. Fixation or near fixation of the tolerance allele was reached in 21% of simulations. This suggests that for the parameter space we sampled, fixation of the tolerance allele in both patches is uncommon but not impossible; likewise, there is a roughly 50% chance that the tolerance allele disappears from the population. Allele frequencies in each patch were tightly coupled, with a rank correlation close to 1. Because of this tight coupling, both patches can not reach their optimal equilibrium; migration load must appear in at least one. Dispersal between patches is one plausible mechanism for this tight coupling; movement of birds between patches creates gene flow, known to homogenize allele frequencies in spatially structured populations.

Steady state allele frequencies in both patches were most sensitive to fledging probabilities of the maladaptive alleles, which is a measure of selection pressure. Both patches had stronger response to the fledging probability of tolerant birds in the uncontaminated patch (a positive relationship in both patches), although because of overlapping confidence intervals we can not definitively conclude that selection pressure in one patch has a stronger effect. This confirms our first prediction that local selection pressure has a strong influence on allele frequencies. However, our results also suggest local allele frequencies are sensitive to selection pressures

in distant patches. Stated differently, local selection pressure has global effect. This supports the above observation that allele frequencies in the two patches are tightly coupled; local selection pressure strongly affects allele frequencies locally but also in neighboring patches.

Local selection pressure also had a strong, significant effect on reproductive output in each patch. In both patches, selection pressure in the neighboring patch had a much stronger effect than local selection pressure. This result is counterintuitive; one would expect that local selection pressures should control fitness, not distant ones. This may further support the conclusion that the allele frequencies are very tightly coupled by gene flow. Another surprising result is that strength of selection pressure has a different relationship with fitness in each patch; lower selection pressure in the contaminated patch correlates with higher local fitness, but lower selection pressure in the uncontaminated patch leads to lower fitness. Perhaps this is due to the fact that in half of trials, the contaminated patch still had near fixation of the wild type allele, and thus an increase in fledging probability for wild type birds would increase overall fitness. The relationship between maladaptive fledging probability and the weighted total fitness was positive, and stronger than all other correlations other than life history traits. This suggests that selection pressure is the strongest influence on reproductive output in a two-patch landscape.

Dispersal likelihood α also had significant partial rank correlation with tolerance allele frequencies in both patches. This relationship was negative in both patches, suggesting that as barriers to dispersal increase, frequency of the tolerance allele decreases in both patches. We expected with our third prediction that increasing α would decrease tolerance in the contaminated patch but increase it in the contaminated patch; thus our prediction was only partially supported. The ι parameter did not have a significant effect on allele frequencies, meaning that the rate of nest overturning and vigor of competition for nests likely does not affect allele frequencies. Our prediction about the relationship between α and within patch fitnesses was also only partially supported, as increasing barriers to dispersal increased fitness in the uncontaminated patch but not the contaminated patch. This suggests that the mechanism that motivated our third prediction, easier dispersal between patches introducing maladaptation and reducing fitness, is not evident in our model. Existing literature shows instances where increasing gene flow increases population fitness by adding needed variation to the gene pool and avoiding inbreeding depression, but this explanation does not work for model, as we examine only a single allele and have no inbreeding depression effects in our model.

The only life history trait that has a statistically significant influence on allele frequencies is the fecundity of subadults, a positive relationship. There is no

intuitive reason for why fecundity of subadults would have the greatest effect on allele frequencies; perhaps a life-cycle sensitivity analysis could provide insight into the crucial role of subadult birth in population dynamics [8]. Additionally, density dependence d had no significant effect on allele frequency, population size, or gene flow in any patch, although it did have a significant effect on net gene flow into the contaminated patch. Gene flow is sensitive to life history parameters, although in several cases the relationship is negative. This is the opposite relationship of the one expected in our fourth prediction, suggesting that gene flow decreases as fecundity or survivorship probability decreases. Perhaps this is because higher fecundity or survivorship does not directly lead to a higher floater population; this would make sense if selection pressure was so strong that created open nests at a rate faster than could be filled by reproducing birds. However, the fact that there is still non-zero gene flow at equilibrium in many trials suggests that there is still movement between populations, indicating that there is a class of floaters and that all nests in at least one patch are full. This question needs more consideration.

Gene flow did correlate with allele frequencies, although the strength of correlation was much stronger in the uncontaminated patch than in the contaminated one. Correlation with fitness was also present, and made intuitive sense: increases in gene flow mean an influx of maladaptive alleles. This validates our second prediction, although only weakly. Asymmetrical migration has been referenced in several studies as a threat to viability and a homogenizing force [4][5][21][39]. We found that asymmetrical gene flow did have a strong significant relationship with allele frequencies and fitnesses in each patch. This supports our fifth prediction that asymmetrical migration into a patch increases maladaptation in that patch. However, the mechanism for creating asymmetrical gene flow is not easily evident: the only parameter that gene flow asymmetry is sensitive to is density dependence. If there was evidence that more floaters led to more gene flow, then a density dependence relationship could provide a logical mechanism for creating gene flow asymmetry through source-sink dynamics. For example, larger values of d in a source patch could mean more floaters which then occupy open territories in the sink habitat. However, since this model found no direct tie between floaters and population growth and gene flow, we are unable to draw that conclusion at this time. Surprisingly, the size of the uncontaminated population ϕ_u did not have a significant correlation on net gene flow. Patch size alone, then, can not determine the steady state of populations and allele frequencies in the whole population.

Although not as strong as the coupling of allele frequencies, population sizes showed a strong correlation. In nearly all instances, populations in at least one patch fell below analytical steady state, suggesting population reduction are inevitable with the introduction of a contaminant. However, complete population

crashes were rare, representing only 5% of all trials. In some of these instances of population crash, reproductive output was above 1, suggesting that $R' < 1$ for a patch (or the whole population) is not a sufficient condition for population crash. This suggests that there may be unseen source-sink dynamics in the model.

4.2 Future Work

Future work on this model could benefit from more biologically relevant data. Some data from zebra finches were used for selection pressures, but other data points were arbitrarily chosen, as were the ranges of parameters tested in latin hypercube sampling. It is possible that lack of variation in output exists because the parameter range sampled was not wide enough to capture different dynamics. Field work with attention devoted to a single species could provide a reliable dataset for future analysis and application of the model.

This model observes only steady-state dynamics of contaminated populations. This means that valuable information about population dynamics while the population is changing are lost. For example, observations of a small set of randomly chosen trials showed that in all instances observed, the contaminated patch subpopulation experienced a very sharp dip in the first time steps of the model, then slowly climbed back to its new steady state. Information about the minimum population size shown could be valuable, especially for future studies that account for genetic drift and stochastic effects which could bring a subpopulation to extinction. Information about the time until fixation could also be of importance for handling legacy contaminants, like those in the Shenandoah. Knowing how far into the process of recovery a population is may allow one to assess how close to a new steady state a population is.

Incorporation of more aspects of source-sink dynamics would improve this study; source-sink analysis has already been applied to mercury contamination of frogs in the Shenandoah [40]. Categorizing populations as sources or sinks requires some parameter or process with a threshold value above which a population is a source or below which a population is a sink. It would make sense that R' acts as a sufficient parameter for this, but the fact that a subpopulation can crash in some instances where $R' > 1$ implies that another parameter, perhaps somehow related to reproductive output, could be used for this threshold.

Finally, more complicated landscapes can be translated into patch networks and studied in depth. We saw homogenization of allele frequencies and, to a lesser extent, populations in a simple two-patch case. However, even the addition of a third patch can introduce can introduce different landscape topologies; studies of processes such as dispersal and adaptation have been studied on three-patch land-

scapes and found that even at this simple level, different patch configurations can lead to different outcomes [34] [42]. Including more patches and then varying landscape configurations would create another interesting input variable for sensitivity analysis.

4.3 Conclusion

In a model of spread of an allele for tolerance to an environmental contaminant, allele frequency and reproductive output were observed to correlate most with strength of selection and asymmetry of migration. These correlations support hypotheses and prior research on causes of migration load and population differentiation. Other, weaker correlations were also observed, but did not support our predictions, and more research is needed to validate their significance and tie them back to biological mechanisms. Several other predictions about relationships between parameters and response variables were not observed. In a two-patch simulation, allele frequencies in each patch were tightly coupled, suggesting there is some homogenizing process preventing divergence of allele frequencies. Population sizes of each patch also showed coupling, although not as tight. This demonstrates that local dynamics and processes can have strong effects on distant subpopulations.

References

[1] Akcakaya, H.R., Atwood, J.L. 1997. A Habitat-Based Metapopulation Model of the California Gnatcatcher. Conservation Biology. 11 (2): 422-434.

[2] Araki, H., Berejikian, B.A., Ford, M.J., and Blouin, M.S. 2008. Fitness of hatchery-reared salmonids in the wild. Evolutionary Applications 1 (2): 342-355.

[3] Bolnick, D.I., Nosil, P. 2007. Natural Selection in Populations Subject to a Migration Load. Evolution. 61 (9): 2229-2243.

[4] Bolnick, D.I., Caldera, E.J., Matthews, B. 2008. Evidence for asymmetric migration load in a pair of ecologically divergent stickleback populations. Biological Journal of the Linnean Society. 94 (2): 273-287.

[5] Bouchy, P. Theodorou, K., Couvet, D. 2005. Metapopulation viability: influence of migration. Conservation Genetics 6: 75-85.

[6] Bourne, E.C., Bocedi, G., Travis, J.M.J., Pakeman, R.J., Brooker, R.W., & Schiffers, K. 2014. Between migration load and evolutionary rescue: dispersal, adaptation and the response of spatially structured populations to environmental change. Proceedings of the Royal Society of Biology. 281

[7] Brown, J.L. 1969. The Buffer Effect and Productivity in Tit Populations. The American Naturalist 103 (932): 347-354.

[8] Caswell, H. 2001. Matrix Population Models: Construction, Analysis and Interpretation. Sinauer: Sunderlad, MA, USA.

[9] Coulon, A., Fitzpatrick, J.W., Bowman, R., Lovette, I.J. 2012. Mind the gap: genetic distance increases with habitat gap size in Florida scrub jays. Biology letters of the Royal Society 8, 582-585.

[10] Cristol, D.A., Brasso, R.L., Condon, A.M., Fovargue, R.E., Friedman, S. L., Hallinger, K.K., Monroe, A.P. & White, A.E. 2008. The movement of aquatic mercury through terrestrial food webs. Science 320:335. Aquatic Toxicology

[11] Dias, P.C. Blondel, J. 1996. Local Specialization and Maladaptation in the Mediterranean Blue Tit (Parus caeruleus). Oecologia 107 (1): 79-86.

[12] DelVecchio, R., Friedman, S., Unsworth, R. 2010. South River and South Fork of the Shenandoah River Natural Resource Damage Assessment: Draft Damage Assessment Plan. Prepared for the United States Department of the Interior, Fish and Wildlife Service, Commonwealth of Virginia Department of Environmental Quality.

[13] Durell, S.E.V., Carke, R.T. 2004. The buffer effect of non-breeding birds and the timing of farmland bird declines. Biological Conservation 120 (3): 375-382.

[14] Esler, D. 2000. Applying Metapopulation Theory to Conservation of Migratory Birds. Conservation Biology 14(2): 366-372.

[15] Fitzpatrick, S.W., Gerberich, J.C., Kronenberger, J.A., Angeloni, L.M., Funk, W.C. 2015. Locally adapted traits maintained in the face of high gene flow. Ecology letters 18: 37-47.

[16] Gyllenberg, M., Metz, J.A.J. 2001. On fitness in structured metapopulations. Journal of Mathematical Biology 43: 545-560.

[17] Hanski, I. 1998. Metapopulation Dynamics. Nature. 396:41-49.

[18] Hanski, I., Gilpin, M. 1991. Metapopulation dynamics: brief history and conceptual domain. Biological Journal of the Linnean Society. 42: 3-16.

[19] Harris, R.J., Reed, J.M. 2002. Behavioral barriers to non-migratory movements of birds. Ann. Zool. Fennici 39: 275-290.

[20] Kendall, M.G. 1942. Partial Rank Correlation. Biometrica 32 (3/4): 277-283.

[21] Lenormand, T. 2002. Gene flow and the limits to natural selection. Trends in Ecology and Evolution. 17 (4): 183-189.

[22] Logan, J.D. 2009. A Primer on Population Genetics. Department of Mathematics, University of Nebraska, Lincoln; Lincoln, NE. 14 - 17

[23] Lopez, S., Rousset, F., Shaw, F.H., Shaw, R.G., Ronce, O. 2008. Migration load in plants: role of pollen and seed dispersal in heterogeneous landscapes. Journal of Evolutionary Biology.21 294:309.

[24] Marino, S., Hogue, I.B., Ray, C.J., Kirschner, D.E. 2008. A methodology for performing global uncertainty and sensitivity analysis in systems biolo-Kendall42gy. Journal of Theoretical Biology 254: 178-196.

[25] Meyer, J.N., Di Giulio, R.T. 2002. Patterns of heritability of decreased EROD activity and resistance to PCB 126-induced teratogenesis in laboratory-reared offspring of killifish (Fundulus heteroclitus) from a creosote-contaminated site in the Elizabeth River, VA, USA. Marine Environmental Research 54: 624 - 626.

[26] Meyer, J.N., Smith, J.D., Winston, G.W., Di Giulio, R.T. 2003. Antioxidant defenses in killifish (Fundulus heteroclitus) exposed to contaminated sediments and model prooxidants: short-term and heritable responses. Aquatic Toxicology 65(4): 377 - 395.

[27] Nosil, P. 2009. Adaptive Population Divergence in Cryptic Color-Pattern Following in a Reduction of Gene Flow. Evolution. 63(7): 1902-1912.

[28] Opdam, P. 1991. Metapopulation theory and habitat fragmentation: a review of holarctic breeding bird studies. Landscape Ecology 5(2): 93-106.

[29] Padilla, B.J., Rodewald, A.D. 2014. Avian metapopulation dynamics in a fragmented urbanizing landscape. Urban Ecosystems 17: 239-250.

[30] Pujol, G., Iooss, B., Janon, A. 2015. sensitivity: Sensitivity Analysis. R package version 1.11. http://CRAN.R-project.org/package=sensitivity.

[31] Pulliam, H.R. 1988. Sources, Sinks and Population Regulation. The American Naturalist. 132(5): 652-661.

[32] Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S. 2008. Global Sensitivity Analysis: The Primer. Wiley & Sons Ltd: Chichester, West Sussex, England.

[33] Scheiman, D.M., Dunning, J.B., With, K.A. 2007. Metapopulation Dynamics of Bobolinks Occupying Agricultural Grasslands in the Midwestern United States. American Midland Nauralist 158(2): 415-423.

[34] Spromberg, J., John, B.M., Landis, W.G. 1998. Metapopulation Dynamics: Indirect Effects and Multiple Distinct Outcomes in Ecological Risk Assessment. Environmental Technology and Chemistry 17 (8): 1640 - 1649.

[35] Theodorakis, C.W. 2001. Integration of Genotoxic and Population Genetic Endpoints in Biomonitoring and Risk Assessment. Ecotoxicology 10: 245-256.

[36] Turchin, P. 2003. Complex Population Dynamics. Princeton University Press: Princeton, NJ, USA.

[37] Varian-Ramos, C.W., Swaddle, J.P., Cristol, D.A. 2013. Familial Differences in the effects of mercury on reproduction in zebra finches. Environmental Pollution 182: 316-323.

[38] Varian-Ramos, C.W., Swaddle, J.P., Cristol, D.A. 2014. Mercury reduces avian reproductive success and imposes selection: an experimental study with adult- or lifetime-exposure in zebra finch. PLoS ONE 9(4): e95674.

[39] Vuilleumier, S., Possingham, H.P. 2006. Does colonization asymmetry matter in metapopulations? Proc. R. Soc. B 273: 1637-1642.

[40] Willson, J.D., Hopkins, W. A. 2013 Evalutating the Effects of Anthropogenic Stressors on Source-Sink Dynamics in Pond-Breeding Amphibians. Conservation Biology 27 (3): 595-604

[41] Woltmannm, S., Kreiser, B.R., Sherry, T.W. 2012. Fine-scale genetic population structure of an understory rainforest bird in Costa Rica. Conservation Genetics 13: 925-935.

[42] Wu, J., Vankat, J.L., and Barlas, Y.L. 1993. Effects of patch connectivity and arrangement on animal metapopulation dynamics: a simulation study. Ecological Modelling, 65: 21-254.